

## Milestone 2: Design document

Nicolas Binuani, Rodolphe Calvet, Sully Dilou, Delin JIA, Zakaria TIZAF

### I. Reminder

Reminder on the topic, the datasets used, where the data come from, and their formats.

**Dataset:** Health\_Europe\_dataset\_somediseases

**Source:**

[http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth\\_cd\\_anr&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_cd_anr&lang=en)

**Format :** csv

- A description of the problem
  - Who are the users ?
    - OMS - Doctors - Pharmacy companies who want to follow the evolution of some disease depending on age, country or gender and apply or modify health plans.
  - What are their backgrounds ?
    - Health domain expertise and maybe some politicians
  - What are they trying to understand from the data?
    - The evolution of some diseases in each country and during a decade to validate or not some health policies taken (example with ALCOVE European program for Alzheimer disease 2011)
  - Is your visualization aimed primarily at *exploring* or *communicating* the data?
    - The first goal is to communicate on the evolution of some diseases in Europe.
    - After that, we can also use it to explore some new insights and change health European programs.
- A description of the data
  - What are the characteristics of the data, the attributes, the size of the dataset, etc.

The dataset gives the total number of deaths by disease, gender, age, and year from 2001 to 2010 in Europe.

It contains 190000 rows and 6 columns.

Year	GEO	SEX	AGE	Disease	Number
2001	Belgium	Female	From 25 to 29 years	Pneumonia	261
...	...	...	...	...	...

Columns Time (Year) and Value (Number) are integers, the other columns present textual data.

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 179718 entries, 0 to 190079
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  -
0   TIME                  179718 non-null  int64
1   GEO                   179718 non-null  object
2   UNIT                  179718 non-null  object
3   SEX                   179718 non-null  object
4   AGE                   179718 non-null  object
5   ICD10                 179718 non-null  object
6   Value                 179718 non-null  int64
7   Flag and Footnotes    4752 non-null    object
dtypes: int64(2), object(6)
memory usage: 12.3+ MB
```

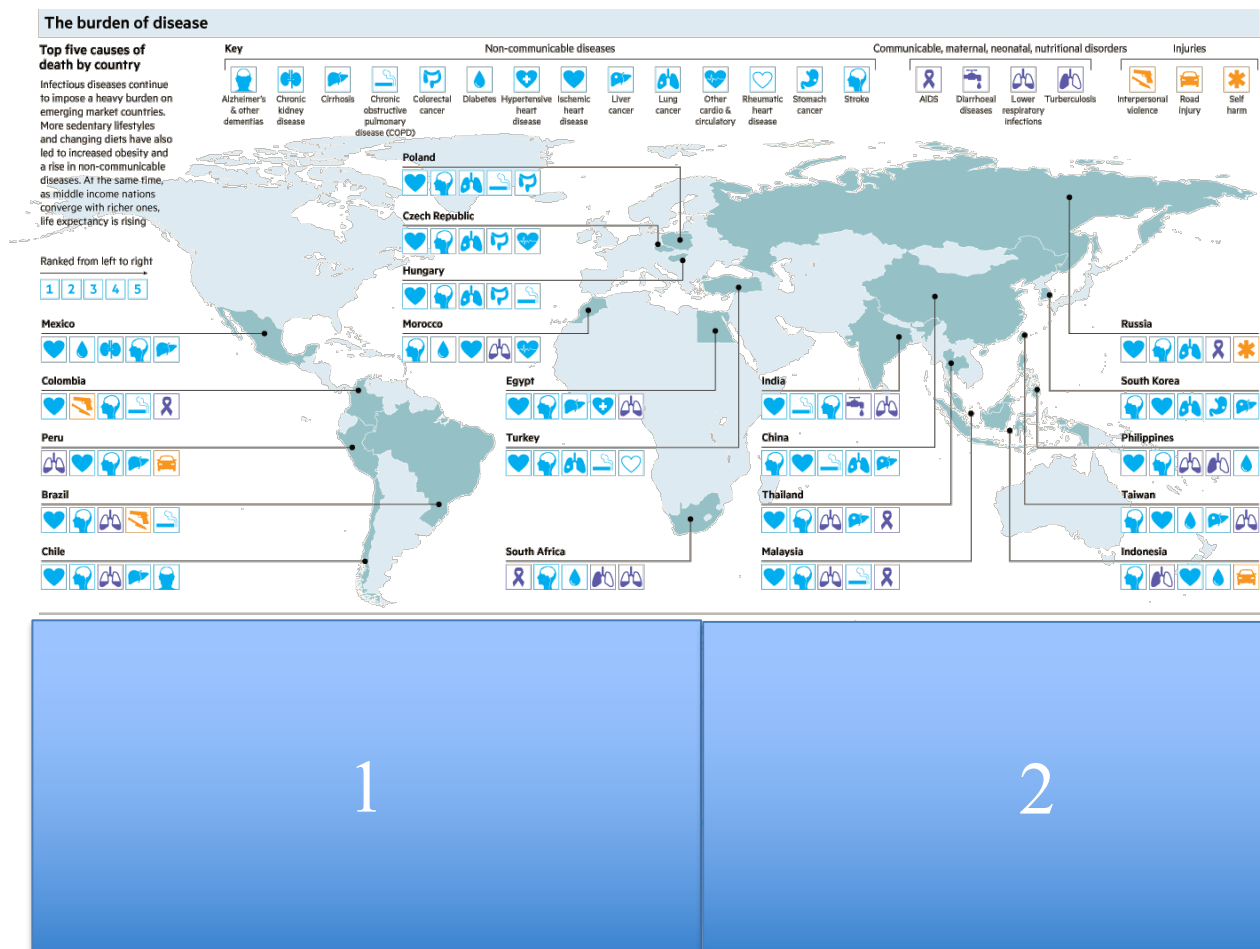
## II. Our three first design ideas

### II.1 First approach

Our goal on this project is to have a tool for health experts, doctors, and politicians about evolution of some most known disease.

In fact, we thought about CovidTracker application that is a little like a kind of dashboard to follow disease evolution.

This is a kind of representation that we want to use:



We have a map which we can select different countries and focus on some disease available in our database.

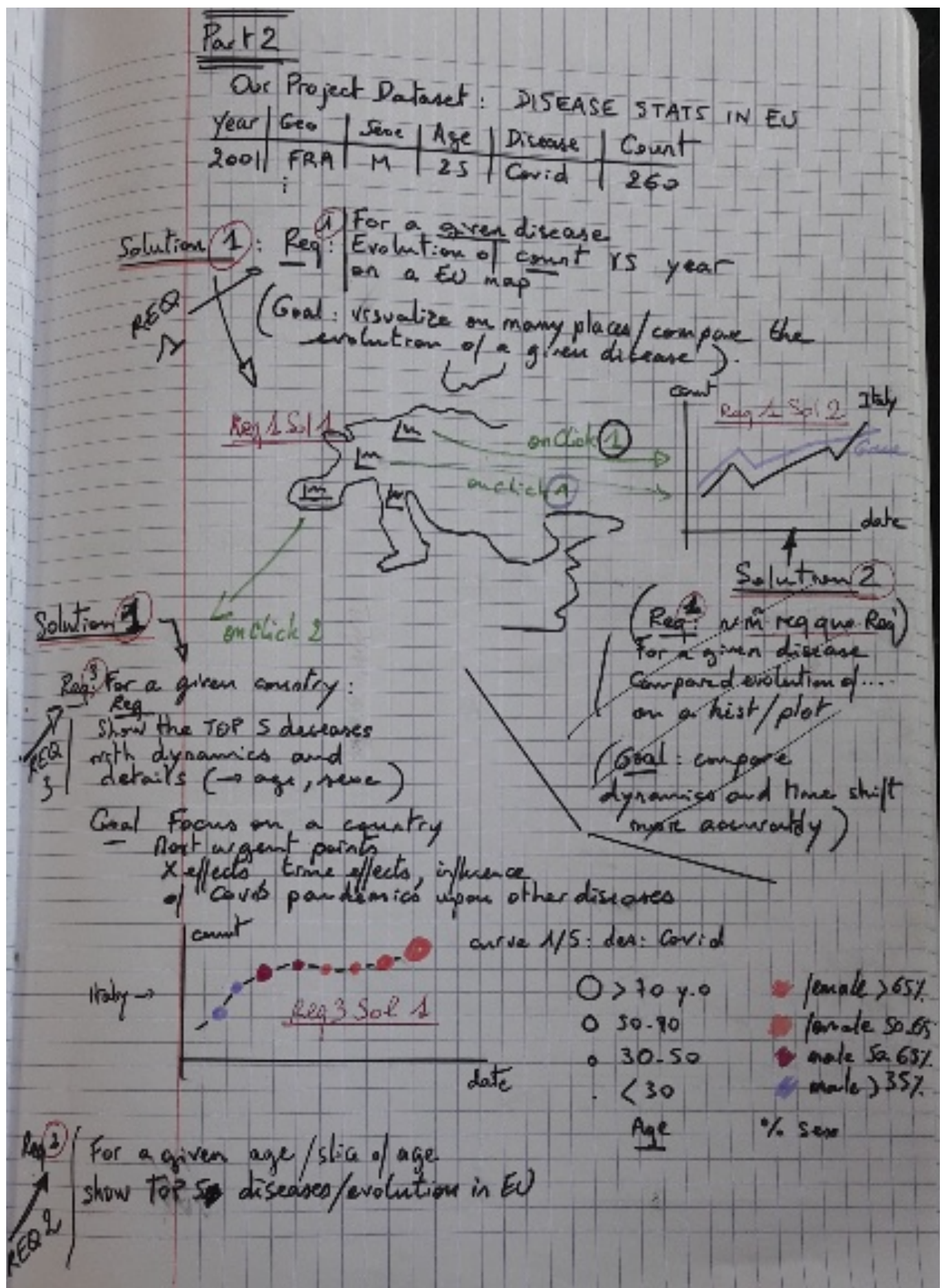
After we have 2 or 3 different plots to give more information like a percentage barplot depending on different variables.

- We can also have a barplot for one disease chosen and have a kind of ranking with a barplot / or a rectangular diagram board to focus on which country is more affected (refer to part II.3)
- We can have gender like the example, and we can focus on age repartition depending on disease observed and for each country. (Refer to part II.3)

To go deeper we have drawn some other features that we may add on our dashboard visualization.

## II.2 Design map

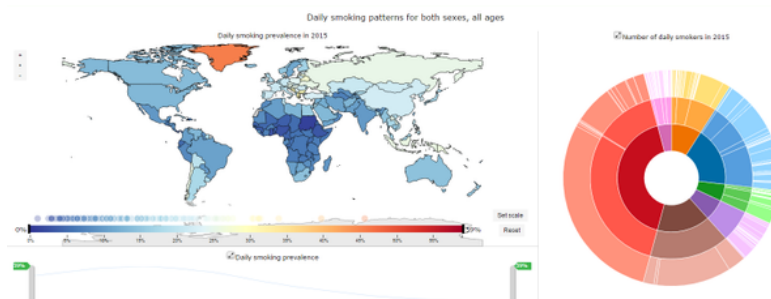
In first case we think about a Europe map to represent disease.



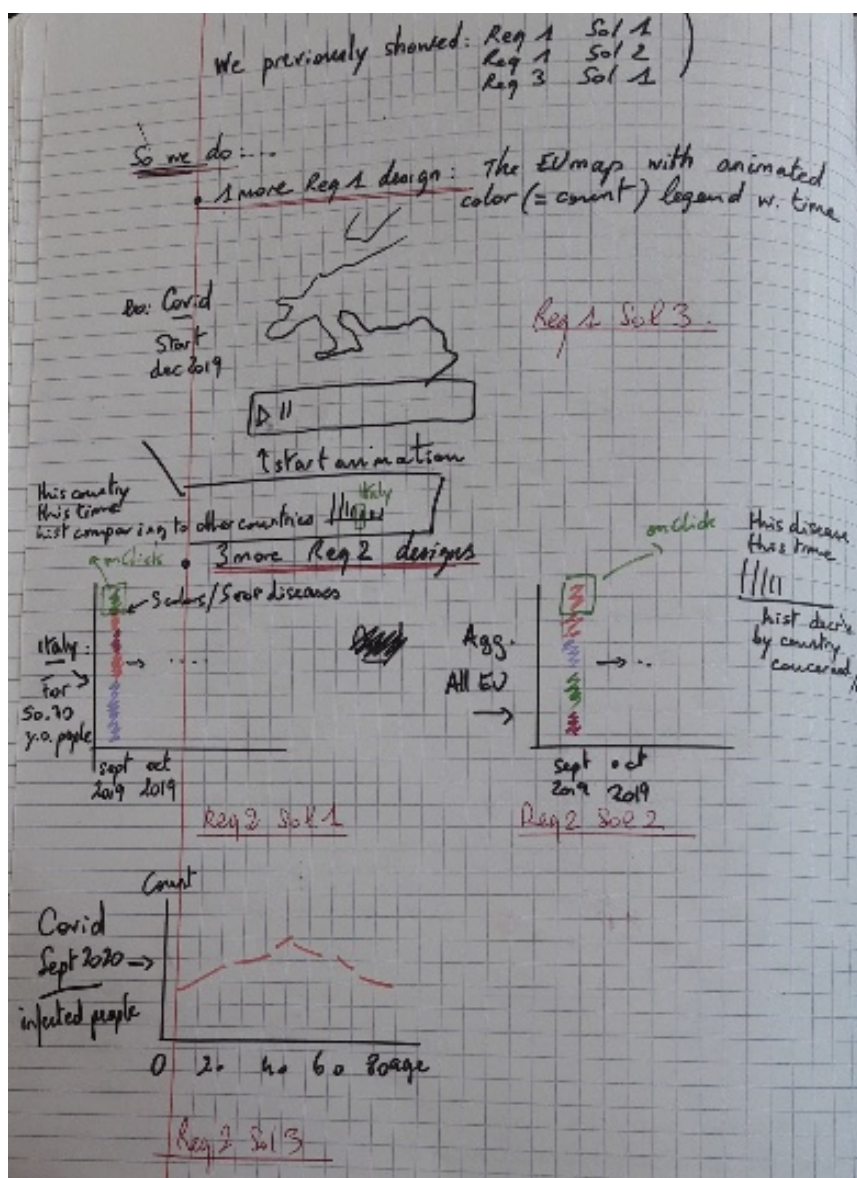


We also add some features such as:

- We can select one country on a map and have a time serie plot of one disease also selected and for each year point size depending on age patient.
- We can select one country on a map and have evolution of count victims compare to mean of Europe indeed we have a kind of threshold
- Also, one point we think about is maybe to select one country on a map and have a kind of country disease pie chart such as this:



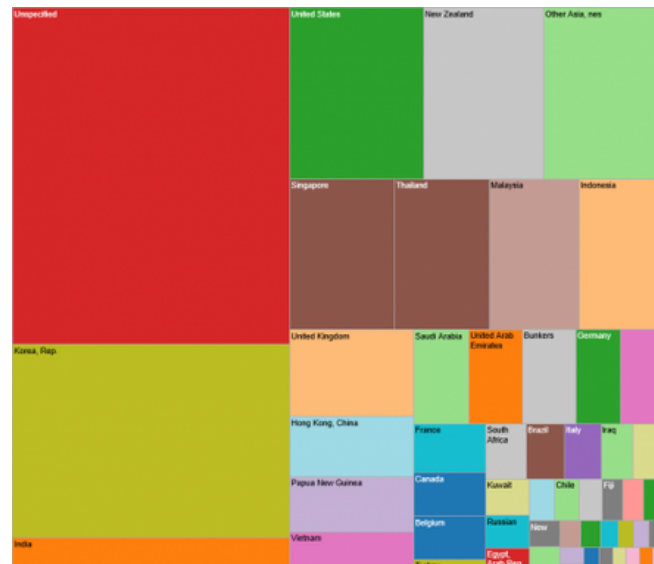
In a second approach, we think about add animated color legend like representing with Request1 Solution 3 just below:



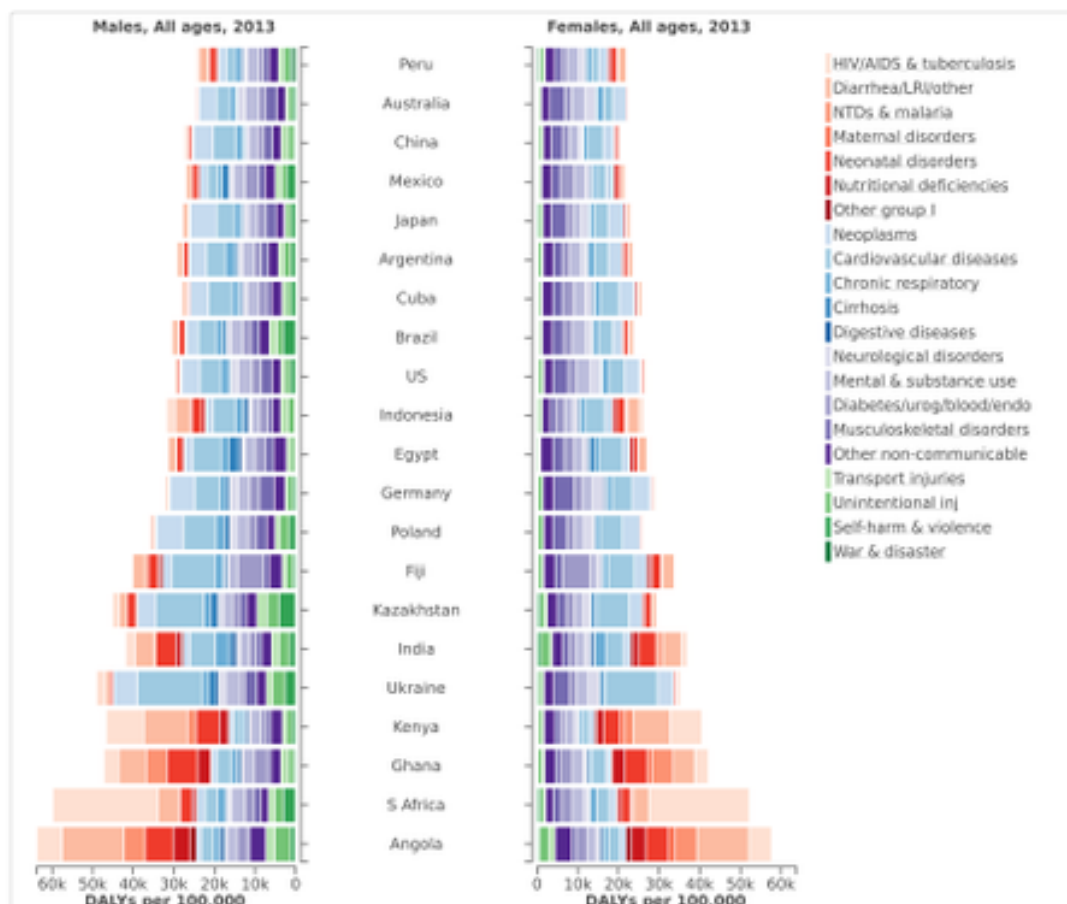
## II.3 Dashboard plots

Kind of plots that can be integrated below the map to have more insights directly visible. We need for this to add a radio button with all disease available in our database.

Other way to represent it, like this (plot 1 from dashboard in part II.1) :



Also one visualization that can be directly present in the dashboard is a barplot depending on gender with a percentage age scale and for one disease selected: (plot 2 from dashboard in part II.1)



Here is a kind of summary of which request and how we want it to interact with our visualizations.

• 2 more Req 3 designs

Req 3 sol 1 en multiplot (# countries)

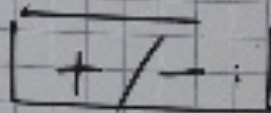
Req 3 Sol 2

Req 3 sol 1 avec

points plottés en  
X: qd le pays a + de malades  
que moy EU

Req 3 Sol 3

O qd — — malades



~~Design 1:~~

Req 1: - carte EU → positioning plots on  
it will hide / be difficult  
to / me time  
+ carte EU → shows geo dynamics  
of diseases

Req 2

+ dynamics in time due to x scale  
= time  
+ onClick → comparison  
country / EU

Req 3

- takes time to read information  
due to symbols entropy ++  
+ all features of data on  
one graph (for one country  
though).